PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/00, G01N 33/68, A61K 38/16

A1

(11) International Publication Number: **WO-98/16638**

(43) International Publication Date: 23 April 1998 (23.04.98)

(21) International Application Number: PCT/US96/16495

(22) International Filing Date: 16 October 1996 (16.10.96)

(71) Applicants: THE ROCKEFELLER UNIVERSITY [US/US]; 1230 York Avenue, New York, NY 10021 (US). REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; Morrill Hall, 100 Church Street, S.E., Minneapolis, MN 55455 (US).

(72) Inventors: COWBURN, David; 215 Dudley Avenue, Westfield, NJ 07090 (US). ZHENG, Jie; Apartment 2D, 515 E. 78th Street, New York, NY 10021 (US). BARANY, George; 1813 Prior Avenue North, Falcon Heights, MN 55113 (US). XU, Qinhong; 1626 Eustic Street #4B, Lauderdale, MN 55108 (US).

(74) Agents: JACKSON, David, A. et al.; Klauber & Jackson, 411 Hackensack Avenue, Hackensack, NJ 07601 (US).

(81) Designated States: AU, CA, JP, MX, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: CONSOLIDATED LIGANDS WITH INCREASED AFFINITIES AND METHODS OF USE THEREOF

(57) Abstract

The invention relates to consolidated ligands, which include amino acid sequences which bind to at least two domains of a target protein. The consolidated ligands have increased affinity for the target protein over ligands directed to a single domain. Particularly preferred ligands include those directed to the Src Homology 2 and 3 (SH2, SH3) domains of eukaryotic protein tyrosine kinases, which are involved in complex regulation of the enzymic activity, obligatory for differentiation and growth control. Ligands of this kind of design may be widely useful as reagents in the investigation of SH interactions, and as leads for design of therapeutic agents.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

CONSOLIDATED LIGANDS WITH INCREASED AFFINITIES AND METHODS OF USE THEREOF

GOVERNMENT SUPPORT

The present invention was made with funding under NIH grants GM-47021, GM-42722 and 51628. The government may have certain rights in the invention.

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to rationally designed consolidated ligands, which are synthesized by linking at least two linear peptide ligands together in an orientation which allows for binding of the consolidated ligand to its substrate with a greater affinity than either linear peptide alone. The consolidated ligands of the invention are useful for the investigation of ligand-substrate interactions, particularly those interactions involved in signal transduction pathways, and as diagnostic and therapeutic agents.

Description of the Related Art

Eukaryotic cellular signal-transduction pathways that are initiated by transmembrane receptors with associated tyrosine kinases rely on two small protein modules, known as SH2 and SH3 domains (src homology domains), for mediating many of the protein-protein interactions that are necessary for transmission of the signal [Cantley et al (1991) *Cell* **64**:281-302; Schlessinger et al (1992) *Neuron* **9**:383-391; Pawson et al (1993) *Curr. Biol.* **3**:434-442]. These domains were first discovered in cytoplasmic (non-receptor) protein tyrosine kinases such as the src oncogene product, thus leading to the term 'src homology domains' [Sadowski et al (1986) *Mol. Cell. Biol.* **6**:4396-4408].

The unique importance of these domains became clear with the discovery of the crk oncogene product, which consists of little more than an SH2 and an SH3 domain fused to the viral gag protein, but is capable of transforming cells [Mayer et al (1988) *Nature* **332**:272-275]. SH2 and SH3 domains have been identified in molecules with distinct functions that act downstream from the receptors for, among others, epidermal growth

actor (EGF), platelet-derived growth factor (PDGF), insulin and interferon, and the T-cell receptor [Koch et al (1991) *Science;* 252:668-674].

The key aspect of the function of SH2 and SH3 domains is their ability to recognize particular amino acid sequences in their target proteins: SH2 domains bind tightly to phosphorylated tyrosine residues [Anderrson et al (1990) *Science;* 250:979-982; Matsuda et al (1990) *Science* 248:1537-1539; Moran et al (1990) *Proc. Natl. Acad. Sci. USA* 87:8622-8626; Mayer et al (1991) *Proc. Natl. Acad. Sci. USA;* 88:627-631; Songyang et al (1993) *Cell* 72:767-778] whereas SH3 domains bind to unmodified peptide sequence that are rich in proline and hydrophobic amino acids [Cicchetti et al (1992) *Science* 257:803-806; Ren et al (1993) *Science* 259:1157-1161]. The modular nature of these domains is made clear by the fact that they occur in different positions in the polypeptide chains of the intact proteins of which they are a part, and that the binding functions can often be reproduced by isolated domains. Although SH2 and SH3 domains frequently occur close together in sequence, some proteins have only one or the other domain, and some have more that one version of either domain. Proteins that contain more than one of these domains do not always maintain a strict spacing or particular order between the domains.

Proteins that contain SH2 and SH3 domains fall into two broad classes: those with catalytic functions and those without. Enzymes that contain these domains include cytoplasmic tyrosine kinases (e.g., Src, Abl, Lck), phosphotyrosine phosphatases, phospholipase C gamma (PLC gamma), *ras* GTPase-activating protein (GAP) and nucleotide exchange factors [Koch et al (1991)]. The SH2 and SH3 domains in these proteins serve to modulate enzyme activity, or to target the enzymes to certain cellular locations. The specific mechanism of the regulation of enzymes by SH2 and SH3 domains is poorly understood. The modulation may simply be a consequence of the recruitment of the protein in question via SH2 or SH3 interactions to a specific region of the cell, such as near the plasma membrane, or to a specific multi-protein complex, such as those associated with the cytoplasmic regions of transmembrane receptors. Modulation might also occur by simple blocking or unblocking of sites, as suggested for the c-Src tyrosine kinase [Roussel et al (1991) *Proc. Natl. Acad. Sci. USA* 88:10696-10700; Lu et al (1993) *Oncogene* 8:1119-1126].

The second class of SH2 and SH3 containing proteins do not exhibit enzymatic activity. Some of these "adapter" proteins bind to enzymes that contain the appropriate targeting sequences (i.e., phosphorylated tyrosines or specific proline-rich sequences) and

modulate their activity. For example, the p85 subunit of phosphatidyl inositol 3-kinase (PI3K) contains SH3 and SH2 domains, but has no catalytic activity. Instead, it serves to potentiate the activity of the kinase subunit, and is also important in localizing this enzyme to the vicinity of activated receptors [Cantley et al (1991); Panay et al (1992) *EMBO* **11**:461-472; Carpenter et al (1993) *J. Biol. Chem.* **268**:9478-9483; Shoelson et al (1993) *EMBO* **12**:795-802]. Adapter proteins are the focus of considerable attention today because of their recently discovered role in linking tyrosine-phosphorylated receptors to *ras* signaling pathway (for example, see [Rozakis-Adgock et al (1992) *Nature* **360**:689-692; Egan et al (1993) *Nature* **363**:45-51; Oliver et al (1993) *Cell* **73**:179-191; Simon et al (1993) *Cell* **73**:169-177; Skolnik et al (1993) *Science* **260**:1953-1955]. Adapter proteins can also serve as transcription factors [Fu et al (1993) *Cell* **74**:1135-1145].

The structural studies on SH2 and SH3 domains demonstrate that these are indeed compactly folding units that can function when removed from their natural contexts. Each of the domains contains within it well defined binding sites for specific interactions with their respective targets. The peptide-binding surfaces of both domains are distal to the termini of the polypeptide chain, consistent with the fact that these domains function in many different structural contexts. There is however evidence that SH2 and SH3 domains may interact with each other in the much larger signal-transduction molecules of which they are a part. For example, the SH3 domain of Src appears to be required for effective inhibitory interaction between the phosphorylated Src tail and the SH2 domain. Intriguingly, an example of a structure with SH2- and SH3-like modules has been found in an unexpected place; the crystal structure of the bacterial DNA-binding protein contains a DNA-binding domain followed by an enzymatic domain that is closely related in its architecture to the SH2 domain, with substrate being bound at a site that is analogous to the phosphotyrosine-binding site. An SH3 like domain of unknown function follows. There is an elaborate interfacial surface between the two domains, and this structure may provide a view of one of the ways in which these structural modules can be packed against each other.

Overduin et al [*Cell* 70:697-704 (1992)], which is hereby incorporated by reference in its entirety, determined the structure of the *abl* SH2 product, a protein of 109 residues and 12.1 kDa, by multidimensional nuclear magnetic resonance spectroscopy. It is a compact spherical domain with a pair of three-stranded antiparallel beta sheets and a C-terminal alpha helix enclosing the hydrophobic core. Three arginines project from a short N-terminal alpha helix and one beta sheet into the putative phosphotyrosine

binding site, which lies on a face distal from the termini. Comparison with other SH2 sequences supports a common global fold and mode of phosphotyrosine binding for this family.

Waksman et al [*Nature* 358:646-653 (1992)], which is hereby incorporated by reference in its entirety, determined the crystal structure of the SH2 domain of v-*src* complexed with tyrosine-phosphorylated peptides. They determined the structure to contain a central antiparallel beta-sheet flanked by two alpha-helices, with peptide binding mediated by the sheet, intervening loops and one of the helices. The specific recognition of phosphotyrosine involves amino-aromatic interactions between lysine and arginine side chains and the ring system in addition to hydrogen-bonding interactions with the phosphate.

Ren et al [(1993) *Science* 259:1157-1161], which is hereby incorporated by reference in its entirety, identified a consensus sequence for the SH3 binding sites of various proteins, which is

X P X X P P P Ψ X P (SEQ ID NO:1).

There is absolute conservation of the three proline sites, which are identified as being critical for binding.

Brugge [(1993) *Science* 260:918-919] highlighted the features of the SH2-phosphotyrosine interactions which make them especially attractive for structure-based drug design, as follows: (1) phosphopeptides as small as five amino acids can interfere with these binding interactions *in vitro*, and due to the small size of the binding sites and the fact that they are derived from a sequence of contiguous amino acids, the design of peptide mimetics is facilitated; (2) peptide mimetic inhibitors will compete effectively for the natural ligand because the rate of dissociation is high; (3) SH2 domains, when expressed as soluble independent domains, readily form crystals and thus, their structure can be determined; and (4) SH2 domains are small enough to be analyzed by high-resolution solution NMR, facilitating the structural analyses of the interaction between the SH2 domain and the peptide mimetics.

Zhu et al [(1994) *Oncogene* 9:1379-1385], which is hereby incorporated by reference in its entirety, studied the ability of an abl SH2 glutathione-S-transferase (GST) fusion protein to bind the phosphorylated epidermal growth factor (EGF), platelet derived growth factor (PDGF) and insulin receptors. They determined that the region surrounding the tyrosine residue at position 1086 was critical for binding. Substitution

of this tyrosine for phenylalanine in the EGF receptor reduced the binding affinity for the abl SH2 domain 10-fold, and a phosphorylated peptide modeled on the sequence surrounding this tyrosine specifically inhibited abl SH2 binding.

Activation of transcription factors may involve dimerization through SH2-phosphotyrosyl peptide interactions [Shuai et al (1994) *Cell* 76:821-828]. Interaction of the SH2 and SH3 domains may be required to effect regulation of proteins which are tyrosine phosphorylated [Eck et al (1994) *Nature* 368:764-769].

Songyang et al [(1995) *Nature* 373:536-539] reported that cytosolic tyrosine kinases preferentially phosphorylate peptides recognized by their own SH2 domains or closely related SH2 domains, while receptor tyrosine kinases preferentially phosphorylate peptides recognized by subsets of group III SH2 domains. Thus, a point mutation in one type of tyrosine kinase can shift its peptide substrate specificity.

The crystal structures of SH3 domains from spectin (X-ray) [Musacchio et al (1992) *Nature* 359:851-855] *src* (NMR) [Yu et al (1992) *Science* 258:1665-1668], p85(NMR) [Booker et al (1993) *Cell* 73:813-822; Koyama et al (1993) *Cell* 72:945-952], PLC*γ*(NMR) [Kohda et al (1993) *Cell* 72:953-960] and the tyrosine kinase *fyn* (X-ray) [Noble et al (1993) *EMBO* 12:2617-2624] have also been reported. The peptide-binding site of the SH3 domain has been identified by analyzing chemical shift changes induced by the addition of peptides [Yu et al (1992); Booker et al (1993)].

U.S. Patent No. 5,352,660 to Pawson discloses a method for assaying a medium for the presence of a substance that affects an SH2-phosphorylated ligand regulatory system. Illustrative SH2-containing proteins disclosed therein included Fps, Src, Abl, GAP, PLC*γ*, v-Crk, Nck, p85 P13K, Tensin and Vav. The SH2 domains mediate the interactions of phosphorylated ligands with signalling proteins which regulate pathways that control gene expression, cell division, cytoskeletal architecture and cell metabolism. Substances which block SH2-containing oncoproteins were hypothesized to be useful in the treatment of thyroid, breast carcinoma, stomach cancer and neuroblastoma, chromic myelogenous leukemia and acute lymphocytic leukemia.

U.S. Patent No. 5,369,008 to Arlinghaus et al, discloses methods for detecting the presence of a BCR-ABL or abnormal ABL gene product using a monoclonal antibody specific for the SH2 region of the ABL gene product.

Abl is a nonreceptor tyrosine kinase originally identified as a viral (v-Abl) product from the transforming gene of Abelson murine leukemia virus [Goff et al (1980) *Cell* 22:777-785]. In human leukemias, cellular Abl is involved by chromosomal translocation containing N-terminal BCR. Simple overexpression of Abl in cultured cells does not lead to transformation, but structurally altered forms do. Abl contains several well-defined sequential segments, and some of these other than the SH domains have been studied in detail [Van Etten et al (1994) *J. Cell Biol.* 124:325-340; Wang (1994) *TIBS* 19:373-376]. The normal function of Abl is probably modification of expression during the cell cycle, possibly mediated by the binding of the retinoblastoma (RB) protein [Wang (1994)].

From experiments with *abl* the evidence is briefly that i) mutation/deletion of SH3 activates the kinase activity/transforming activity of the normally nontransforming proto-oncogene products, [Mayer and Baltimore (1994) *Mol. Cell. Biol.* 14:2883-94], while ii) mutation/deletion of SH2 decreases or eliminates the kinase activity/transforming activity of transforming versions of the kinase (i.e., those that already have some kind of activating mutation) [Muller et al (1993) *Proc. Natl. Acad. Sci. USA* 90:3457-3461], iii) substitution of SH2s still is permissive of transformation, with a different pattern of phosphorylated products. In the absence of both SH2 and SH3, the proteins produced appear to be non-transforming. As an additional complication, there are reports of mutations in SH2 being able to activate the kinase activity/transforming activity of *src*. A possibility is that these arise by structural mutations which reduce the SH3 regulation of SH2. All of the genetic evidence is so far consistent with SH2 function being required for the activated, transforming state, and with SH3 serving to negatively regulate SH2 function in the normal protein. Abl is clearly subtly different from the Src family of kinases, e.g., Src and Lck, in which a 'cis' inactivation mechanism involving the intramolecular association of pY site with an SH2 in the same chain, with possible mediation by the same-chain SH3 e.g., [Superti-Furga et al (1993) *EMBO J.* 12:2625-2634, and references therein]. The current hypothesis for the Abl mechanism of down regulation in normal cells is that there are SH2, and SH3 interactions close to the kinase, involving a weak interaction with another unidentified soluble factor, presumably of low affinity but relatively high specificity.

In order to identify possibly physiologically significant proteins binding to Abl, extensive studies using the yeast two-hybrid system were undertaken [Fields and Song (1989) *Nature* 340:245-246]. A specific site for the direct interaction between Crk-I SH3 and Abl was identified, just C-terminal to the kinase domain of Abl, both *in vitro* and in

mammalian cells [(Ren and Baltimore (1994) *Genes and Development*, and Conference on Signal Transduction, NY Acad. Med, 9/94). This is the "AB3-1" site indicated in the above scheme of *abl*. Crk-I is phosphorylated by Abl kinase on binding. Simultaneously, the Hanafusa laboratory showed that Crk-II is phosphorylated by Abl kinase (Feller et al (1994) *EMBO J.* 13:2341-2351). Crk-I and Crk-II are alternatively spliced products and contain an SH2 and SH3 domain. Crk-II has a second C-terminal SH3 domain. Crk-I-SH3 binds to Abl *in vitro*. Similar experiments establish other adapter proteins, Grb2, and NCk as *in vitro* binders to Abl at the AB3-2 and AB3-3 sites (see above).

Such large "mosaic" proteins present formidable problems in general to structural analysis, in that, for example, for many of the tyrosyl kinases, and for other SH3- and SH2-containing proteins, the separate segments clearly have different and defined biological functions. An important general question is then the degree to which order and position of the various motifs are significant and effect function, and how to investigate this.

If the various sites merely are required to contribute individual functions, such as bringing the components of a large complex together, then interdomain interactions may be minimal. If very precise sequential and three-dimensional positioning are needed, it is currently difficult to see how that arises since SH2 and SH3 domains appear to be well folded entities, and there is little evidence for required interactions. Very recently, Eck and Harrison [Eck et al (1994) *Nature* 368:764-769) have shown that in the crystal form, a specific dimer contact is formed in Lck SH(32) construct, when the SH2 is occupied by a ligand. No evidence for dimerization in Lck SH(32) is available in solution, and there are minimal contacts between same-chain SH3 and SH2 segments. This last point is confirmed in Abl SH(32). Significant amounts of biological evidence point to a role for positional specific interactions between the SH3 and the SH2 of Src (Superti-Furga, 1993). There is then an unresolved issue concerning the apparent effects of SH(32) interactions in transformation, and current structural information which is ambiguous about such interactions. The Eck/Harrison model of control of kinase activity by homomolecular dimerization has been widely adopted, despite the lack of evidence for in-solution dimerization. A separate proof of modulation of activity by homomolecular dimerization involving SH2's comes from the work reported here later [Shuai et al (1994) *Cell* 76:821-828], for STAT.

Protein dynamics are generally considered to play major roles in enzymic catalysis,

ligand recognition, and allosteric effects [Wagner (1993) *Curr. Op. Stru. Biol.* 3:748-754]. NMR methods address different time scales of motion, from the relatively fast, using spin relaxation measurements, through analysis of shifts, nOes, coupling etc., in the medium range, to the slow observation of chemical exchange from intensity of chemically shift separated species (e.g., HD amide exchange). Fast internal dynamics measured via ^{15}N relaxation, and to a lesser extent to date from ^{13}C relaxation, provide a picture of the motions of the backbone and sidechains respectively, and in many cases this information complements the density of defined nOes in this portion of the molecules and is similar to the B-factors in crystal structure determinations (Powers et al (1993) *J. Magn. Reson. Ser.* **B101**:325-327). Significant technical advances in speed of measurement by indirect detection, in reduction of second order effects, in corrections for chemical exchange, in estimation of precision etc., have rendered these methods of general applicability [Skelton et al (1993) *J. Mag. Res.* **B102**:253-264; Kordel et al (1992) *Biochem.* **31**:4856-4866; Fushman et al (1994) *J. Biomol. Struct. & Dynamics* 4:61-78; Schneider et al (1992) *Biochem.* **31**:3645-3652]. Analysis of such motions generally uses the "model-free" formalism [Lipari et all (1982) *J. Am. Chem. Soc.* 104:4546-4559; Clore et al (1990) *J. Am. Chem. Soc.* **112**:4989-4991], which is of general utility, but has so far been only linked in a limited way to specific motions or conformational fluctuations [Fushman et al (1994); Palmer and Case (1992) *J. Am. Chem. Soc.* **114**:9050-9067] or energetics [Akke et al (1993) *J. Am. Chem. Soc.* **115**:9832-9833]. In slower time ranges, variations of vicinal coupling constants may provide a view of dynamic fluctuations, but applicability to larger proteins (cf., peptides (Cowburn et al (1983) *J. Am. Chem. Soc.* 105:7435-7442) is still technically highly limited. Similarly, dynamic analysis of chemical shifts in intermediate conformation exchange is technically limited to a few sites (e.g., some aromatic ring flips). At the slower time ranges, hydrogen exchange from amides and other exchangeable hydrogens, and the accessibilities to the water solvent are powerful tools [Harayama et al (1989) *Biochem.* **28**:4312-4317; Clore and Gronenborn (1992) *J. Mol. Biol.* **223**:853-856]. While there are multiple studies of the perturbations of structure by bound ligands, few reports deal with the dynamics of connected domains. In the case of calmodulin, a very complete picture indicating flexibility of the central helix has been presented, along with a detailed analysis of the likely effects of motional anisotropy [Barbato et al (1992) *Biochem.* **31**:5269-5278].

More generally, the problems of organization and structure of large mosaic proteins need to be attacked to provide the benefits of structural insight to many cellular signalling processes of central importance in human biology. Protein structure determination by

NMR has now reached a stage where it can bear respectable comparison with medium resolution crystal structures of small proteins (<20 kD), and can provide useful and unique, if selective, views of the dynamics of interplay between ligands the various domains binding them, and the domains themselves.

The understanding of the pathways involving ligand-binding domains, particularly SH domains, is complicated by the substantial range of individual specificities of these domains, and their potential combination into large proteins that can form multiple homo- and heteromolecular associations. Moreover, there are technical limitations to current structural approaches. Thus, the interactions of multiple domains are poorly understood. Such interactions are likely to be of significance in explaining more fully the complete activities of the proteins containing these domains, particularly those involved in signal transduction pathways.

SUMMARY OF THE INVENTION

In view of the aforementioned deficiencies attendant with prior art methods of identifying ligands for multiple protein domains, and of structural analysis of ligand binding domains, it is evident that there exists a need in the art for methods of designing such ligands, which can be used to modify the activity of those proteins and to analyze the structure of the ligand-binding domains.

In accordance with the present invention, multidomain complexes termed "consolidated ligands" are provided. These ligands, having multiple binding portions, may be expected to bind with high affinity and specificity when a linker between the two affine segments is of the correct length. The linker may have little affinity, or may be modified specially to interact with the intervening, "passive" structure. The consolidated ligands do not necessarily resemble any natural ligand.

The proposed "consolidated" ligands are similar in concept to affinity reagents, with the modification that the second functionality is a binding element rather than a reactive moiety. Consolidated ligands may also be useful reagents for studies of the cell biology of the signal transduction complexes, such as in their additional combination with antibodies or reporter groups, e.g., fluorescent tags, and may provide leads into possible classes of diagnostic and therapeutic agents in the many areas of pathology in which SH domains are involved. This approach will be useful where micromolar affinities of ligands to single SH domains provide insufficient affinity, and hence, specificity for

pharmacological action (Gibbs et al (1994) *Cell* 77:175-178; Levitzki and Gazit (1995) *Science* 267:1782-88).

The consolidated ligands of the present invention can be prepared by linking at least two ligands for at least two protein domains, which domains may be the same or different. By "domain" is meant any amino acid sequence which represents a fragment of a larger polypeptide, which fragment has a characteristic sequence, structure, and/or activity. For example, a preferred embodiment of the invention is a consolidated ligand containing ligands to the SH2 and SH3 domains of molecules which function downstream from cellular receptors having protein tyrosine kinase (PTK) activity. These domains function in the signal transmission from such receptors to various intracellular sites. The domains to which the ligands are directed may be from the same or different proteins.

The ligands for the domains may be arranged linearly, or in branched form, and may be arranged in the same or opposite directions (i.e., NH_2 to COOH end). The ligands may be of any length such that they have a suitable binding affinity for the domain to which they are directed, however, they are preferably from about 4 amino acids in length to about 25 amino acids in length, and more preferably from about 6 to about 12 amino acids in length. Likewise, the linker(s) connecting the domains may be of any length such that the multiple ligands are spaced and oriented for optimal reactivity with the target protein. The linkers may be designed such that they have no affinity for the target protein, or optionally, they may be modified such that they have an affinity for sequences in the target protein which are intervening between the domains to which the ligands are directed. The linkers may be comprised of any molecule, including peptidomimetic, and non-peptidic linkers, or may be comprised of other subligands including lipids, steroids, carbohydrates, and nucleic acids. The linker is preferably chosen for optimal flexibility and length, which will obviously vary with the choice of ligands, and is preferably about 10-50 Å, more preferably about 20-40 Å, and most preferably about 30-35 Å. In a preferred embodiment of the invention, the linker is comprised of glycyl moieties, preferably about 2 to 20 glycyls, more preferably about 3 to 12 glycyls, and most preferably about 5 to 8 glycyls.

The receptor from which the signals (which are transduced by the signal transducing protein or target protein) originate may bind to any external protein factor, but preferred factors include epidermal growth factor (EGF), erythropoietin (EPO), fibroblast growth factor (FGF, including acidic or aFGF and basic or bFGF and FGF3), insulin-like growth

factor (IGF, including IGF-1 and IGF-2), interleukins (IL, including IL-1, IL-1, IL-3, IL-4, and IL-6), leukemia inhibitory factor (LIF), platelet-derived growth factor (PDGF), stem cell factor (SCF), transforming growth factor (TGF, including TGF-α and TGF-β), tumor necrosis factor (TNF) and vascular endothelial growth factor (VEGF).

In addition, the signal transducing protein which contains the "domain" may be any such signal transducing protein, however preferred embodiments include Abl, APRF, Arg, Bcl-2, Blk, Btk, c-erbB-2, CAK, cbl, Cbl-B, CCK-2, CD25, CDK, CFSR, clk, crk, crkl, E2A, EGFR, ena, erbB2, ERK-2, Esk, FAK, Fes, Fgr, Fit3, Flk2, FLT1, fos, Fps, FUS3, Fyn, gap, grb2, grb7, HOX11, HER2, HGF/SF, IGR-IR, IGFR, int-2, iyk, Jak3, jun, KDR, KIT, Lck, Lyl1, lyn, MAP kinase, Matk, MCK-10, MET, MO15, mta1, MUC1, nck, nek, NEU, p53, PDGFR, PI3K, PIM, PLC-gamma, PTB1b, PTP1C, PTPN6, RAF, Rak, ras, RB, Ret, Rlk, ROS, SCL, Shc, Sos, SRC, STK1, SYK, Syp, Tec, tie, TRK, tsk, Ttg1, Ttg2, Txk, Vav, and yes.

Because of the extensive cellular and biological studies of Abl, and the preexisting solution structural studies published and in progress, it is an exceptionally suitable system, having both interesting biological implications and protein structural features, to pursue studies of the effects of ligation, interdomain interaction, and domain dynamics.

The pathologies for which the present consolidated ligands may be used in treatment include any human, animal or plant pathology which is associated with an increase or decrease in the activity of a protein or factor containing a biologically active domain, especially those proteins or factors described above. Pathologies or cellular states which are particularly well suited to such treatment include agammaglobulinemia, AIDS, ALL, angiogenesis, breast cancer, carcinoma, chromic myelogenous leukemia, colon carcinoma, colorectal cancer, diabetes, erythroleukemia, gastric cancer, hematopoiesis, Kaposi's Sarcoma, leukemia, liver regeneration, Lyme disease, megakaryocytopoiesis, melanoma, neuroblastoma, organogenesis, osteopetrosis, ovarian hyperstimulation syndrome, placental development, severe combined immunodeficiency, ulcerative colitis and Wilms tumor.

The concept of the consolidated ligand contemplates that specific factors exist for correspondingly specific ligands, such as for the SH2 and SH3 domains and the like, as described earlier. Accordingly, the exact structure of each consolidated ligand will understandably vary so as to achieve this ligand and activity specificity, as well as preferably achieving a higher affinity than a corresponding naturally occurring ligand. It

is this specificity and the direct involvement of the protein containing the ligand-binding domain in the chain of events leading to gene activation and/or cytoplasmic activities, that offers the promise of a broad spectrum of diagnostic and therapeutic utilities.

The present invention naturally contemplates several means for preparation of the consolidated ligand, including as illustrated herein automated peptide synthesis and known recombinant techniques, and the invention is accordingly intended to cover such synthetic and recombinant preparations within its scope. The isolation of the cDNA sequences encoding the amino acid sequences of the consolidated ligands disclosed herein facilitates the reproduction of the consolidated ligands by such recombinant techniques, and accordingly, the invention extends to expression vectors prepared from these DNA sequences for expression in host systems by recombinant DNA techniques, and to the resulting transformed hosts.

The invention includes an assay system for screening of potential drugs effective to modulate the activity of the protein containing the domain to which the ligand binds by interrupting or potentiating its biological activity. In one instance, the test drug could be administered to a cellular sample with the ligand that activates the protein or factor, or an extract containing the activated protein or factor, to determine its effect upon the binding activity of the ligand to any chemical sample (including DNA), or to the test drug, by comparison with a control.

The assay system could more importantly be adapted to identify drugs or other entities that are capable of increasing the binding affinity of the ligand to the proteins or factors, either in the cytoplasm or in the nucleus, thereby inhibiting or potentiating the activity of the protein or factor. Such assay would be useful in the development of drugs that would be specific against particular cellular activity, or that would potentiate such activity, in time or in level of activity. For example, such drugs might be used to modulate cellular response to shock, or to treat other pathologies.

In yet a further embodiment, the invention contemplates consolidated ligands which act as antagonists of the activity of a protein or factor. In a specific embodiment, the antagonist can be a peptide having a sequence which binds to a portion of a portion of an SH2 or SH3 domain of a protein tyrosine kinase. In a specific example, *infra*, such peptides are shown to be capable of disrupting the signal transducing activity of *src* and *src*-like proteins.

One of the characteristics of the present consolidated ligands is that they may have enhanced affinity for the domain of the protein or factor to which they bind. Another characteristic of the consolidated ligands is that they may have enhanced specificity for the domain of the protein or factor to which they bind.

The diagnostic utility of the present invention extends to the use of the present consolidated ligand in assays to screen for proteins or factors containing the domain or domains to which the consolidated ligand binds.

The present invention likewise extends to the development of antibodies against the consolidated ligand(s) and of consolidated ligand(s)/target, including naturally raised and recombinantly prepared antibodies. For example, the antibodies could be used to screen expression libraries to obtain a clone which encodes the consolidated ligand(s). Such antibodies could include both polyclonal and monoclonal antibodies prepared by known genetic techniques, as well as bi-specific (chimeric) antibodies, and antibodies including other functionalities making them well suited for additional diagnostic use conjunctive with their capability of modulating the protein or factor's activity.

In particular, consolidated ligands for specifically phosphorylated factors can be selected and are included within the scope of the present invention for their particular ability in following activated protein. Thus, activity of the protein or factor or of the specific polypeptides believed to be causally connected thereto may therefore be followed directly by the assay techniques discussed later on, through the use of an appropriately labeled quantity of the consolidated ligand.

Thus, the consolidated ligands, their analogs and/or analogs, and any antagonists or antibodies that may be raised thereto, are capable of use in connection with various diagnostic techniques, including immunoassays, such as a radioimmunoassay, or using for example, a consolidated ligand that has been labeled by either radioactive addition, or radioiodination.

In an immunoassay, a control quantity of the consolidated ligand, antagonists or antibodies thereto, or the like may be prepared and labeled with an enzyme, a specific binding partner and/or a radioactive element, and may then be introduced into a cellular sample. After the labeled material or its binding partner(s) has had an opportunity to react with sites within the sample, the resulting mass may be examined by known techniques, which may vary with the nature of the label attached. For example,

consolidated ligands for specifically phosphorylated factors may be selected and appropriately employed in the exemplary assay protocol, for the purpose of following activated protein as described above.

In the instance where a radioactive label, such as the isotopes ^{3}H, ^{14}C, ^{32}P, ^{35}S, ^{36}Cl, ^{51}Cr, ^{57}Co, ^{58}Co, ^{59}Fe, ^{90}Y, ^{125}I, ^{131}I, and ^{186}Re are used, known currently available counting procedures may be utilized. In the instance where the label is an enzyme, detection may be accomplished by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, other spectroscopic, amperometric or gasometric techniques known in the art.

The present invention includes an assay system which may be prepared in the form of a test kit for the quantitative analysis of the extent of the presence of the protein or factor to which the consolidated ligand binds, or to identify drugs or other agents that may mimic or block their activity. The system or test kit may comprise a labeled component prepared by one of the radioactive and/or enzymatic techniques discussed herein, coupling a label to the consolidated ligands, their agonists and/or antagonists, and one or more additional immunochemical reagents, at least one of which is a free or immobilized ligand, capable either of binding with the labeled component, its binding partner, one of the components to be determined or their binding partner(s).

In a further embodiment, the present invention relates to certain therapeutic methods which would be based upon the binding activity of the consolidated ligand(s), its (or their) subunits, or active fragments thereof, or upon agents or other drugs determined to possess the same activity. A first therapeutic method is associated with the prevention of the manifestations of conditions causally related to or following from the activity (whether abnormally increased or decreased) of the protein or factor to which the consolidated ligand binds, binding activity of the ligand or its subunits, and comprises administering a consolidated ligand capable of modulating the activity of the protein or factor to which it binds, or subunits thereof, either individually or in mixture with each other in an amount effective to prevent the development of those conditions in the host. Also, the blockade of the action of specific tyrosine phosphatases in the dephosphorylation of activated (phosphorylated) proteins or factors presents a method for potentiating the activity of the consolidated ligand.

More specifically, the therapeutic method generally referred to herein could include the method for the treatment of various pathologies or other cellular dysfunctions and

derangements by the administration of pharmaceutical compositions that may comprise the consolidated ligands which act as effective inhibitors or enhancers of activation of the protein or factor containing the domain to which the consolidated ligand binds, or its subunits, or other equally effective drugs developed for instance by a drug screening assay prepared and used in accordance with a further aspect of the present invention. For example, consolidated ligands for the Abelson protein kinase, as represented by SEQ ID NOS:10-16, may be administered to inhibit chromosomal translocation, as in the potentiation of anti-cancer drugs in cancer therapy. Also, the blockade of the action of specific tyrosine phosphatases in the dephosphorylation of activated proteins involved in signal transduction presents a method for potentiating the activity of the signal transducing protein that would concomitantly potentiate other therapies.

In particular, the present consolidated ligands whose sequences are presented in SEQ ID NOS:10-17 herein, their antibodies, agonists, antagonists, or active fragments thereof, could be prepared in pharmaceutical formulations for administration in instances wherein interferon therapy is appropriate, such as to treat chronic viral hepatitis, hairy cell leukemia, and for use of interferon in adjuvant therapy. The specificity of the ligands hereof would make it possible to better manage the after effect of current cancer therapy.

Accordingly, it is a principal object of the present invention to provide a consolidated ligand in purified form that exhibits certain characteristics and activities associated with binding activity to proteins, particularly those proteins involved in signal transduction, and by example those containing SH2 and SH3 domains.

It is a further object of the present invention to provide antibodies to the consolidated ligand, and methods for their preparation, including recombinant means.

It is a further object of the present invention to provide a method for detecting the presence of the proteins which bind the consolidated ligand and their subunits in which invasive, spontaneous, or idiopathic pathological states are suspected to be present.

It is a further object of the present invention to provide a method and associated assay system for screening substances such as drugs, agents and the like, potentially effective in either mimicking the activity or combating the adverse effects of the signal transducing protein and/or its subunits.

It is a still further object of the present invention to provide a method of treatment to control the amount or activity of the signal transducing protein or subunits thereof, so as to alter the adverse consequences of such presence or activity, or where beneficial, to enhance such activity.

It is a still further object of the present invention to provide a method of treatment to control the amount or activity of the target or signal transducing protein or its subunits, so as to treat or avert the adverse consequences of invasive, spontaneous or idiopathic pathological states.

It is a still further object of the present invention to provide pharmaceutical compositions for use in therapeutic methods which comprise or are based upon the consolidated ligand, its subunits, their binding partner(s), or upon agents or drugs that control the production, or that mimic or antagonize the activities of the target protein.

Other objects and advantages will become apparent to those skilled in the art from a review of the ensuing description which proceeds with reference to the following illustrative drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1. Model of the orientations of SH2 and SH3 ligands in a consolidated ligand (I) on the Abl SH(32) protein. Coordinates of individual SH2 and SH3 domains (Overduin et al (1992) *Cell* 70:697-704) were fitted to the approximate orientation of Lck SH(32) in the observed crystal structure (Eck et al (1994) *Nature* 368:764-69). The peptide (I) was created and aligned to the expected binding positions based on modelling of SH2 ligands to Src SH2 (Waksman et al (1993) *Cell* 12:779-90) and SH3 ligands to Abl (Musacchio et al (1994) *Nature Structural Biology* 1:546-51). The positioning of glycyl and branched lysyl residues is arbitrary, other than the restrictions of the subligand positioning. Modelling and display used the Biosym system on an SGI Iris.

FIGURE 2. Representative fluorescence titration data, using the methods described in the legend to Table 1. Titrations are for the branched consolidated ligand, IV, with the three following constructs: SH(32) red, SH3 green, and SH2 blue.

DETAILED DESCRIPTION

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook et al, "Molecular Cloning: A Laboratory Manual" (1989); "Current Protocols in Molecular Biology" Volumes I-III [Ausubel, R. M., ed. (1994)]; "Cell Biology: A Laboratory Handbook" Volumes I-III [J. E. Celis, ed. (1994))]; "Current Protocols in Immunology" Volumes I-III [Coligan, J. E., ed. (1994)]; "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)]; "Transcription And Translation" [B.D. Hames & S.J. Higgins, eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)]; "Immobilized Cells And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

Therefore, if appearing herein, the following terms shall have the definitions set out below.

The terms "signal transducing protein," "protein involved in signal transduction," "SH2 and/or SH3-containing protein," "target protein," and any variants not specifically listed, may be used herein interchangeably, and as used throughout the present application and claims refer to proteinaceous material including single or multiple proteins, and extends to those proteins having the amino acid sequence data and profile of activities previously described and set forth herein and in the Claims. Accordingly, consolidated ligands directed to those proteins, and proteins displaying substantially equivalent or altered activity are likewise contemplated. These modifications, which may be made in the target protein, or the consolidated ligand directed thereto, may be deliberate, for example, such as modifications obtained through site-directed mutagenesis, or may be accidental, such as those obtained through mutations in hosts that are producers of the complex or its named subunits. Also, the terms "signal transducing protein," "protein involved in signal transduction," "SH2 and/or SH3-containing protein," and "target protein," and the consolidated ligands thereto, are intended to include within their scope proteins specifically recited herein as well as all substantially homologous analogs and allelic variations. Such proteins may contain one or more of SH2, and/or SH3 and/or similar domains.

The amino acid residues described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid

residue, as long as the desired functional properties of binding (e.g., to immunoglobulins) is retained by the polypeptide or peptide mimetic. NH2 refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In keeping with standard polypeptide nomenclature, *J. Biol. Chem.*, 243:3552-59 (1969), abbreviations for amino acid residues are shown in the following Table of Correspondence:

TABLE OF CORRESPONDENCE

SYMBOL		AMINO ACID	
1-Letter		3-Letter	
Y		Tyr	tyrosine
G		Gly	glycine
F		Phe	phenylalanine
M		Met	methionine
A		Ala	alanine
S		Ser	serine
I		Ile	isoleucine
L		Leu	leucine
T		Thr	threonine
V		Val	valine
P		Pro	proline
K		Lys	lysine
H		His	histidine
Q		Gln	glutamine
E		Glu	glutamic acid
W		Trp	tryptophan
R		Arg	arginine
D		Asp	aspartic acid
N		Asn	asparagine
C		Cys	cysteine

It should be noted that all amino-acid residue sequences are represented herein by formulae whose left and right orientation is in the conventional direction of amino-terminus to carboxy-terminus. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino-acid residues. The above Table is presented to correlate the three-letter and one-letter notations which may appear alternately herein.

The phrase "pharmaceutically acceptable" refers to molecular entities and compositions that are physiologically tolerable and do not typically produce an allergic or similar untoward reaction, such as gastric upset, dizziness and the like, when administered.

The phrase "therapeutically effective amount" is used herein to mean an amount sufficient to prevent, and preferably reduce by at least about 30 percent, more preferably by at least 50 percent, most preferably by at least 90 percent, a clinically significant change in the S phase activity of a target cellular mass, or other feature of pathology such as for example, elevated blood pressure, fever or white cell count as may attend its presence and activity.

In its primary aspect, the present invention concerns the identification of consolidated ligands, directed to more than a single domain of a particular target protein, which has increased affinity to that target protein over a ligand directed to only one of those domains.

In a particular embodiment, the present invention relates to consolidated ligands directed to target proteins which are proteins involved in signal transduction. In its most preferred aspect, the invention relates to consolidated ligands directed to proteins containing SH2 and/or SH3 domains, and those proteins having both such domains.

As stated above, the present invention also relates to a recombinant DNA molecule or cloned gene, or a degenerate variant thereof, which encodes a consolidated ligand, or a fragment thereof, that possesses an amino acid sequence set forth in SEQ ID NOS:10-17.

The possibilities both diagnostic and therapeutic that are raised by the discovery of the present consolidated ligands, derive from the fact that factors appear to participate in direct and causal protein-protein interaction between a receptor that is occupied by its ligand, and those factors that thereafter directly interface with a gene and effect transcription and accordingly gene activation. As suggested earlier and elaborated further on herein, the present invention contemplates pharmaceutical intervention in the cascade of reactions in which the target protein, or signal transducing protein, is implicated, to modulate the activity initiated by the stimulus bound to the cellular receptor.

Thus, in instances where it is desired to reduce or inhibit the gene activity resulting from

a particular stimulus or factor, an appropriate inhibitor of the target or signal transducing protein could be introduced to block the interaction of the signal transducing protein with those factors causally connected with gene activation. Correspondingly, instances where insufficient gene activation is taking place might be remedied by the introduction of a molecule which binds the consolidated ligand, or its chemical or pharmaceutical cognates, analogs, fragments and the like.

As discussed earlier, the consolidated ligands or their binding partners or other ligands or agents exhibiting either mimicry or antagonism to the consolidated ligands or control over their production, may be prepared in pharmaceutical compositions, with a suitable carrier and at a strength effective for administration by various means to a recipient experiencing an adverse medical condition, particularly those associated with specific cancers or other uncontrolled cell growth, for the treatment thereof. A variety of administrative techniques may be utilized, among them parenteral techniques such as subcutaneous, intravenous and intraperitoneal injections, catheterizations and the like. Average quantities of the consolidated ligands or their subunits may vary and in particular should be based upon the recommendations and prescription of a qualified physician or veterinarian.

Also, antibodies including both polyclonal and monoclonal antibodies, and drugs that themselves bind to the consolidated ligands and/or their subunits may possess certain diagnostic applications and may for example, be utilized for the purpose of detecting and/or measuring the presence of naturally occurring ligands to a target protein. Likewise, an antibody or binding partner to the consolidated ligand may itself possess activities similar to that of the target protein to which the consolidated ligand is directed. For example, the consolidated ligand or its subunits may be used to produce both polyclonal and monoclonal antibodies to themselves in a variety of cellular media, by known techniques such as the hybridoma technique utilizing, for example, fused mouse spleen lymphocytes and myeloma cells. Likewise, small molecules that mimic or antagonize the activity(ies) of the consolidated ligand of the invention may be discovered or synthesized, and may be used in diagnostic and/or therapeutic protocols.

The general methodology for making monoclonal antibodies by hybridomas is well known. Immortal, antibody-producing cell lines can also be created by techniques other than fusion, such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. See, e.g., M. Schreier et al., "Hybridoma Techniques" (1980); Hammerling et al., "Monoclonal Antibodies And T-cell Hybridomas"

(1981); Kennett et al., "Monoclonal Antibodies" (1980); see also U.S. Patent Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,451,570; 4,466,917; 4,472,500; 4,491,632; 4,493,890.

Panels of monoclonal antibodies produced against consolidated ligands can be screened for various properties; i.e., isotype, epitope, affinity, etc. Of particular interest are monoclonal antibodies that either neutralize or mimic the activity of the target protein or its subunits. Such monoclonals can be readily identified in binding assays such as ELISA or Western blot. High affinity antibodies are also useful when immunoaffinity purification of the consolidated ligand is possible.

Preferably, the anti-consolidated ligand antibody used in the diagnostic methods of this invention is an affinity purified polyclonal antibody. More preferably, the antibody is a monoclonal antibody (mAb). In addition, it is preferable for the antibody molecules used herein be in the form of Fab, Fab', $F(ab')_2$ or F(v) portions of whole antibody molecules.

As suggested earlier, the diagnostic method of the present invention comprises examining a cellular sample or medium by means of an assay including an effective amount of a consolidated ligand. As previously discussed, patients capable of benefiting from this method include those suffering from cancer, a pre-cancerous lesion, uncontrolled cell growth, or other like pathological derangement. Methods for isolating the consolidated ligand and inducing antibodies thereto, and for determining and optimizing the ability of consolidated ligands and the antibodies thereto to assist in the examination of the target cells are all well-known in the art.

Methods for producing polyclonal anti-polypeptide antibodies are well-known in the art. See U.S. Patent No. 4,493,795 to Nestor et al. A monoclonal antibody, typically containing Fab and/or $F(ab')_2$ portions of useful antibody molecules, can be prepared using the hybridoma technology described in *Antibodies - A Laboratory Manual*, Harlow and Lane, eds., Cold Spring Harbor Laboratory, New York (1988), which is incorporated herein by reference. Briefly, to form the hybridoma from which the monoclonal antibody composition is produced, a myeloma or other self-perpetuating cell line is fused with lymphocytes obtained from the spleen of an animal hyperimmunized with a consolidated ligand or consolidated ligand-binding portion thereof.

Splenocytes are typically fused with myeloma cells using polyethylene glycol (PEG) 6000. Fused hybrids are selected by their sensitivity to HAT. Hybridomas producing a

monoclonal antibody useful in practicing this invention are identified by their ability to immunoreact with the present consolidated ligand and their ability to modulate specified signal transducing activity in target cells.

A monoclonal antibody useful in practicing the present invention can be produced by initiating a monoclonal hybridoma culture comprising a nutrient medium containing a hybridoma that secretes antibody molecules of the appropriate antigen specificity. The culture is maintained under conditions and for a time period sufficient for the hybridoma to secrete the antibody molecules into the medium. The antibody-containing medium is then collected. The antibody molecules can then be further isolated by well-known techniques.

Media useful for the preparation of these compositions are both well-known in the art and commercially available and include synthetic culture media, inbred mice and the like. An exemplary synthetic medium is Dulbecco's minimal essential medium (DMEM; Dulbecco et al., *Virol.* 8:396 (1959)) supplemented with 4.5 gm/l glucose, 20 mm glutamine, and 20% fetal calf serum. An exemplary inbred mouse strain is the Balb/c.

Methods for producing monoclonal antibodies are also well-known in the art. See Niman et al., *Proc. Natl. Acad. Sci. USA*, 80:4949-4953 (1983). Typically, the present consolidated ligand is used either alone or conjugated to an immunogenic carrier, as the immunogen in the before-described procedure for producing anti-consolidated ligand monoclonal antibodies. The hybridomas are screened for the ability to produce an antibody that immunoreacts with the consolidated ligand or fragment thereof.

The present invention further contemplates therapeutic compositions useful in practicing the therapeutic methods of this invention. A subject therapeutic composition includes, in admixture, a pharmaceutically acceptable excipient (carrier) and one or more of a consolidated ligand, or fragment thereof, as described herein as an active ingredient.

The preparation of therapeutic compositions which contain polypeptides, analogs or active fragments as active ingredients is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions, however, solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified. The active therapeutic ingredient is often mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water,

saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the composition can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents which enhance the effectiveness of the active ingredient.

A consolidated ligand, analog or active fragment can be formulated into the therapeutic composition as neutralized pharmaceutically acceptable salt forms. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide or antibody molecule) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed from the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The therapeutic consolidated ligand, analog- or active fragment-containing compositions are conventionally administered intravenously, as by injection of a unit dose, for example. The term "unit dose" when used in reference to a therapeutic composition of the present invention refers to physically discrete units suitable as unitary dosage for humans, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; i.e., carrier, or vehicle.

The compositions are administered in a manner compatible with the dosage formulation, and in a therapeutically effective amount. The quantity to be administered depends on the subject to be treated, capacity of the subject's immune system to utilize the active ingredient, and degree of inhibition or neutralization of signal transduction or binding capacity desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner and are peculiar to each individual. However, suitable dosages may range from about 0.1 to 20, preferably about 0.5 to about 10, and more preferably one to several, milligrams of active ingredient per kilogram body weight of individual per day and depend on the route of administration. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed by repeated doses at one or more hour intervals by a subsequent injection or other administration. Alternatively, continuous intravenous infusion sufficient to maintain concentrations of ten nanomolar to ten micromolar in the blood are contemplated.

The therapeutic compositions may further include an effective amount of the consolidated ligand or analog thereof, and one or more of the following active ingredients: an antibiotic, a steroid. Exemplary formulations are given below:

Formulations

Intravenous Formulation I

Ingredient	mg/ml
cefotaxime	250.0
consolidated ligand	10.0
dextrose USP	45.0
sodium bisulfite USP	3.2
edetate disodium USP	0.1
water for injection q.s.a.d.	1.0 ml

Intravenous Formulation II

Ingredient	mg/ml
ampicillin	250.0
consolidated ligand	10.0
sodium bisulfite USP	3.2
disodium edetate USP	0.1
water for injection q.s.a.d.	1.0 ml

Intravenous Formulation III

Ingredient	mg/ml
gentamicin (charged as sulfate)	40.0
consolidated ligand	10.0
sodium bisulfite USP	3.2
disodium edetate USP	0.1
water for injection q.s.a.d.	1.0 ml

Intravenous Formulation IV

Ingredient	mg/ml
consolidated ligand	10.0
dextrose USP	45.0
sodium bisulfite USP	3.2
edetate disodium USP	0.1
water for injection q.s.a.d.	1.0 ml

Intravenous Formulation V	
Ingredient	mg/ml
consolidated ligand antagonist	5.0
sodium bisulfite USP	3.2
disodium edetate USP	0.1
water for injection q.s.a.d.	1.0 ml

As used herein, "pg" means picogram, "ng" means nanogram, "ug" or "μg" mean microgram, "mg" means milligram, "ul" or "μl" mean microliter, "ml" means milliliter, "l" means liter.

Another feature of this invention is the expression of the DNA sequences encoding the consolidated ligands disclosed herein. As is well known in the art, DNA sequences may be expressed by operatively linking them to an expression control sequence in an appropriate expression vector and employing that expression vector to transform an appropriate unicellular host.

Such operative linking of a DNA sequence of this invention to an expression control sequence, of course, includes, if not already part of the DNA sequence, the provision of an initiation codon, ATG, in the correct reading frame upstream of the DNA sequence.

A wide variety of host/expression vector combinations may be employed in expressing the DNA sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and Synthetic DNA sequences. Suitable vectors include derivatives of SV40 and known bacterial plasmids, e.g., *E. coli* plasmids col El, pCR1, pBR322, pMB9 and their derivatives, plasmids such as RP4; phage DNAs, e.g., the numerous derivatives of phage λ, e.g., NM989, and other phage DNA, e.g., M13 and Filamentous single stranded phage DNA; yeast plasmids such as the 2μ plasmid or derivatives thereof; vectors useful in eukaryotic cells, such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or other expression control sequences; and the like.

Any of a wide variety of expression control sequences -- sequences that control the expression of a DNA sequence operatively linked to it -- may be used in these vectors to express the DNA sequences of this invention. Such useful expression control sequences include, for example, the early or late promoters of SV40, CMV, vaccinia,

polyoma or adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the *LTR* system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase (e.g., Pho5), the promoters of the yeast α-mating factors, and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof.

A wide variety of unicellular host cells are also useful in expressing the DNA sequences of this invention. These hosts may include well known eukaryotic and prokaryotic hosts, such as strains of *E. coli*, *Pseudomonas*, *Bacillus*, *Streptomyces*, fungi such as yeasts, and animal cells, such as CHO, RI.I, B-W and L-M cells, African Green Monkey kidney cells (e.g., COS 1, COS 7, BSC1, BSC40, and BMT10), insect cells (e.g., Sf9), and human cells and plant cells in tissue culture.

It will be understood that not all vectors, expression control sequences and hosts will function equally well to express the DNA sequences of this invention. Neither will all hosts function equally well with the same expression system. However, one skilled in the art will be able to select the proper vectors, expression control sequences, and hosts without undue experimentation to accomplish the desired expression without departing from the scope of this invention. For example, in selecting a vector, the host must be considered because the vector must function in it. The vector's copy number, the ability to control that copy number, and the expression of any other proteins encoded by the vector, such as antibiotic markers, will also be considered.

In selecting an expression control sequence, a variety of factors will normally be considered. These include, for example, the relative strength of the system, its controllability, and its compatibility with the particular DNA sequence or gene to be expressed, particularly as regards potential secondary structures. Suitable unicellular hosts will be selected by consideration of, e.g., their compatibility with the chosen vector, their secretion characteristics, their ability to fold proteins correctly, and their fermentation requirements, as well as the toxicity to the host of the product encoded by the DNA sequences to be expressed, and the ease of purification of the expression products.

Considering these and other factors a person skilled in the art will be able to construct a variety of vector/expression control sequence/host combinations that will express the

DNA sequences of this invention on fermentation or in large scale animal culture.

It is further intended that consolidated ligand analogs may be prepared from nucleotide sequences of the protein complex/subunit derived within the scope of the present invention. Analogs, such as fragments, may be produced, for example, by pepsin digestion of consolidated ligand material. Other analogs, such as muteins, can be produced by standard site-directed mutagenesis of consolidated ligand coding sequences. Analogs exhibiting binding activity to a target protein such as small molecules, whether functioning as promoters or inhibitors, may be identified by known *in vivo* and/or *in vitro* assays.

As mentioned above, a DNA sequence encoding a consolidated ligand can be prepared synthetically rather than cloned. The DNA sequence can be designed with the appropriate codons for the consolidated ligand amino acid sequence. In general, one will select preferred codons for the intended host if the sequence will be used for expression. The complete sequence is assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge *Nature* 292:756 (1981); Nambair et al., *Science* 223:1299 (1984); Jay et al. *J. Biol. Chem.* 259:6311 (1984).

The invention contemplates development of small molecule analogs of the consolidated ligand sequences, e.g., non-peptidyl molecules that are capable of specifically binding to the multiple domains with high affinity. Such molecules can be designed using rational techniques, based on the detailed structural information provided herein, and by application of available computer assisted structure design.

Another approach uses recombinant bacteriophage to produce large libraries of peptides. Using the "phage method" [Scott and Smith (1990) *Science* 249:386-390; Cwirla et al (1990) *Proc. Natl. Acad. Sci. USA* 87:6378-6382; Devlin et al (1990) *Science* 249:404-406), very large libraries can be constructed (10^6 10^8 chemical entities). A second approach uses primarily chemical methods, of which the Geysen method [Geyson et al (1986) *Molecular Immunology* 23:709-715; Geysen et al (1987) *J. Immunologic Methods* 102:259-274) and the recent method of Fodor et al [(1991) *Science* 251:767-773) are examples. Furka et al [(1988) 14th International Congress of Biochemistry, Volume 5, Abstract FR:013; Furka (1991) *Int. J. Peptide Protein Res.* 37:487-493), Houghton (U.S. Patent No. 4,631,211, issued December, 1986) and Rutter et al (U.S. Patent No. 5,010,175, issued April 23, 1991) describe methods to

produce a mixture of peptides, which may contain L-amino acids, non-peptidyl bonds, and peptidomimetics.

In another aspect, synthetic libraries (Needels et al (1993) *Proc. Natl. Acad. Sci. USA* 90:10700-10074; Lam et al, International Patent Publication No. WO 92/00252, each of which is incorporated herein by reference in its entirety), and the like can be used to screen for analogs of the consolidated ligands according to the present invention.

Screening can be performed on the basis of specific, high affinity binding to a multi-domain construct, inhibition assays, and other binding assays, as described herein.

Preferably, a consolidated ligand of the invention is associated with a carrier molecule to facilitate or enhance biological activity. Preferred carrier molecules facilitate transport through the cell membrane into the cytoplasm. Alternatively, or in addition, the ligand can be associated with a carrier that specifically targets the peptide to a cell in which binding of the ligand is desired. Such targeting molecules include peptides, proteins, antibodies, antibody fragments, lectins, carbohydrates, or steroids. Accordingly, as used herein, the term "targeting molecule" refers to a molecule that can be conjugated to a ligand, which molecule binds specifically to a molecule found *in vivo*, such as a receptor or other recognition molecule or a molecule specific to a cell or cells, etc. Steroids are useful carriers, since they are receptor specific and facilitate transport into the cytoplasm. In a specific embodiment in which the targeting molecule is a peptide, the peptide may contain the well known sequence Arg-Gly-Asp (R-G-D). RGD receptors are found on the surface of cells such as endothelial cells, cancer cells, or ova. Antibodies for use as targeting molecules may be specific for a cell surface antigen. In one embodiment, the antigen is a receptor, for example, an antibody specific for a receptor on a tumor cell can be used. In another embodiment, antibodies specific for leukocyte surface antigens, such as lymphocyte antigens, CD (clusters of differentiation) antigens, and receptors (e.g., T cell antigen receptors) can be conjugated to the peptide. Any antibody known in the art that is specific for a cell antigen can be used as a targeting molecule. In another embodiment, the targeting molecule transferrin may be used. More than one targeting molecule can be used, for example, by using two different molecules to target a peptide to the same *in vivo* location.

A consolidated ligand may be conjugated to the targeting molecule covalently, using well known cross-linking techniques, or by expression of a chimeric construct, e.g., a single chain monoclonal antibody or a transferrin construct containing the ligand

consensus sequence.

Synthetic DNA sequences allow convenient construction of genes which will express consolidated ligand analogs or "muteins". Alternatively, DNA encoding muteins can be made by site-directed mutagenesis of cDNAs encoding the consolidated ligands, and muteins can be made directly using conventional polypeptide synthesis.

A general method for site-specific incorporation of unnatural amino acids into proteins is described in Christopher J. Noren, Spencer J. Anthony-Cahill, Michael C. Griffith, Peter G. Schultz *Science* 244:182-188 (April 1989). This method may be used to create analogs with unnatural amino acids.

The present invention also relates to a variety of diagnostic applications, including methods for detecting the presence of stimuli such as the earlier referenced signal transducing proteins, by reference to their ability to modulate the activities which are mediated by those signal transducing proteins. As mentioned earlier, the consolidated ligand can be used to produce antibodies to itself by a variety of known techniques, and such antibodies could then be isolated and utilized as in tests for the presence of ligands to the target proteins in suspect target cells.

As described in detail above, antibody(ies) to the consolidated ligands can be produced and isolated by standard methods including the well known hybridoma techniques.

The presence of a target protein in cells can be ascertained by the usual binding assays applicable to such determinations. A number of useful procedures are known. Such procedures which are especially useful utilize either the consolidated ligand labeled with a detectable label, an antibody (Ab) to the consolidated ligand labeled with a detectable label, or target protein labelled with an appropriate label. The procedures may be summarized by the following equations wherein the asterisk indicates that the particle is labeled, "T" stands for the target protein and "CL" stands for the consolidated ligand:

A. $T^* + CL = T^*CL$

B. $T + CL^* = TCL^*$

C. $T + CL + Ab^* = TCLAb^*$

The procedures and their application are all familiar to those skilled in the art and accordingly may be utilized within the scope of the present invention. As analogized to immunological procedures, the "competitive" procedure, Procedure A, is described in

U.S. Patent Nos. 3,654,090 and 3,850,752. Procedure C, the "sandwich" procedure, is described in U.S. Patent Nos. RE 31,006 and 4,016,043. Still other procedures are known such as the "double antibody", or "DASP" procedure.

In each instance, the target protein forms complexes with one or more consolidated ligands or binding partners and one member of the complex is labeled with a detectable label. The fact that a complex has formed and, if desired, the amount thereof, can be determined by known methods applicable to the detection of labels.

It will be seen from the above, that a characteristic property of Ab is that it will react with CL. This Ab may be raised in one species, while a second Ab (Ab_2) may be raised in another species with the Ab acting as an antigen to raise the antibody Ab_2. For example, Ab_2 may be raised in goats using rabbit antibodies as antigens. Ab_2 therefore would be anti-rabbit antibody raised in goats. For purposes of this description and claims, Ab will be referred to as a primary or anti-consolidated ligand antibody, and Ab_2 will be referred to as a secondary or anti-Ab antibody.

The labels most commonly employed for these studies are radioactive elements, enzymes, chemicals which fluoresce when exposed to ultraviolet light, and others.

A number of fluorescent materials are known and can be utilized as labels. These include, for example, fluorescein, rhodamine, auramine, Texas Red, AMCA blue and Lucifer Yellow. A particular detecting material is anti-rabbit antibody prepared in goats and conjugated with fluorescein through an isothiocyanate.

The consolidated ligand or its binding partner(s) can also be labeled with a radioactive element or with an enzyme. The radioactive label can be detected by any of the currently available counting procedures. The preferred isotope may be selected from ^{3}H, ^{14}C, ^{32}P, ^{35}S, ^{36}Cl, ^{51}Cr, ^{57}Co, ^{58}Co, ^{59}Fe, ^{90}Y, ^{125}I, ^{131}I, and ^{186}Re.

Enzyme labels are likewise useful, and can be detected by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or gasometric techniques. The enzyme is conjugated to the selected particle by reaction with bridging molecules such as carbodiimides, diisocyanates, glutaraldehyde and the like. Many enzymes which can be used in these procedures are known and can be utilized. The preferred are peroxidase, ß-glucuronidase, ß-D-glucosidase, ß-D-galactosidase, urease, glucose oxidase plus peroxidase and alkaline phosphatase. U.S. Patent Nos.

3,654,090; 3,850,752; and 4,016,043 are referred to by way of example for their disclosure of alternate labeling material and methods.

In a further embodiment of this invention, commercial test kits suitable for use by a medical specialist may be prepared to determine the presence or absence of predetermined signal transducing activity or predetermined signal transducing activity capability in suspected target cells. In accordance with the testing techniques discussed above, one class of such kits will contain at least the labeled consolidated ligand or its binding partner, for instance an antibody specific thereto, and directions, of course, depending upon the method selected, e.g., "competitive", "sandwich", "DASP" and the like. The kits may also contain peripheral reagents such as buffers, stabilizers, etc.

Accordingly, a test kit may be prepared for the demonstration of the presence or capability of cells for predetermined target protein activity, comprising:

(a) a predetermined amount of at least one labeled chemically reactive component obtained by the direct or indirect attachment of the present consolidated ligand or a specific binding partner thereto, to a detectable label;

(b) other reagents; and

(c) directions for use of said kit.

More specifically, the diagnostic test kit may comprise:

(a) a known amount of the consolidated ligand as described above (or a binding partner) generally bound to a solid phase to form a sorbent, or in the alternative, bound to a suitable tag, or plural such end products, etc. (or their binding partners) one of each;

(b) if necessary, other reagents; and

(c) directions for use of said test kit.

In a further variation, the test kit may be prepared and used for the purposes stated above, which operates according to a predetermined protocol (e.g. "competitive", "sandwich", "double antibody", etc.), and comprises:

(a) a labeled component which has been obtained by coupling the consolidated ligand to a detectable label;

(b) one or more immunochemical reagents of which at least one reagent is a ligand or an immobilized ligand, which ligand is selected from the group consisting of:

(i) a ligand capable of binding with the labeled component (a);

(ii) a ligand capable of binding with a binding partner of the labeled component (a);

(iii) a ligand capable of binding with at least one of the component(s) to be determined; and

(iv) a ligand capable of binding with at least one of the binding partners of at least one of the component(s) to be determined; and

(c) directions for the performance of a protocol for the detection and/or determination of one or more components of a binding reaction between the target protein and a specific consolidated ligand thereto.

In accordance with the above, an assay system for screening potential drugs effective to modulate the activity of the target protein/signal transducing protein. The consolidated ligand may be introduced into a test system, and the prospective drug may also be introduced into the resulting cell culture, and the culture thereafter examined to observe any changes in the signal transducing or other activity of the cells, due either to the addition of the prospective drug alone, or due to the effect of added quantities of the known consolidated ligand.

PRELIMINARY CONSIDERATIONS

There have recently been a number of reviews that discuss the function of SH2 and SH3 domains (for example, see [Cantley et al (1991); Schlessinger et al (1992); Pawson et al (1993); Musacchio et al (1992) *FEBS Lett.* 307:55-61; Pawson et al (1992) *Curr. Opin. Struc. Biol.* 2:432-437; Pawson et al (1992) *Cell* 71:559-562; Mayer et al (1993) *Trends Cell Biol.* 318-323; Panayotou et al (1993) *Bioessays* 15:171-177]. One particular recent review describes the specificity of SH2 and SH3 domains and the signal-transduction pathways in which they play a role [Pawson et al (1993)].

Three reports on the structures SH2 domains were published simultaneously. Two of these describe the structures of the SH2 domains of the abl tyrosine kinase [Overduin et al (1992) *Proc. Natl. Acad. Sci. USA* 89:11672-11677; Overduin et al (1992) *Cell* 70:697-704] and the p85 subunit of P13-kinase [Booker et al (1992) *Nature* 338:684-687] determined by NMR in solution and in the absence of ligands. The crystal structure of the src SH2 domain, in complex with two different non-specific phosphopeptides, was also reported at the same time [Wakeman et al (1992) *Nature* 355:646-653]. Subsequent to these papers, two simultaneous reports described the

crystal structures of SH3 domains from spectin (X-ray) [Musacchio et al (1992) *Nature* 359:851-855] src (NMR) [Yu et al (1992) *Science* 258:1665-1668], p85(NMR) [Booker et al (1993) *Cell* 73:813-822; Koyama et al (1993) *Cell* 72:945-952], PLCγ(NMR) [Kohda et al (1993) *Cell* 72:953-960] and the tyrosine kinase *fyn* (X-ray) [Noble et al (1993) *EMBO* 12:2617-2624]. Two of these studies have located the peptide-binding site of the SH3 domain by analyzing chemical shift changes induced by the addition of peptides [Yu et al (1992); Booker et al (1993)].

The architecture of SH2 domains

The structures determined so far make clear that SH2 domains share a common secondary structural scaffolding, with close similarity in tertiary fold. The SH2 domain starts with a very short ß-strand (ßα). An α-helix (αA) follows, which forms part of the phosphotyrosine-binding site. The core of the domain is formed by a continuous ß-meander formed by strands B-E and F (strand F has not been identified as such in p85). These strands form two ß-sheets, which are connected by the long strand (ßD') in the second. Strand F leads into helix aB, which is followed by a loop of variable length (BG). The final ßG is part of the ß sheet formed by strands B-D. A hydrophobic core is formed mainly by the packing of the two helices against the sheets. The peptide binds to a relatively flat surface formed by the two helices an the edge of the central ß-sheet. The amino and carboxyl termini of domain lie close together in space, and on the side of the domain opposite to the peptide-binding surface.

The Src and Lck SH2 domains are closely related, with 53% sequence identity. Abl and p85 are more distantly related to Src, with sequence identities of 37% and 28%, respectively. Despite the somewhat low sequence similarity between some of these domains, the structural elements that make up the core of the SH2 domain superimpose extremely well among Src, Lck,Abl and p85. The positions and relative orientations of the two α-helices and the central ß-sheet are highly conserved. The conformation of the smaller sheet, as well as the EF and BG loops, show considerably more variation, and this variation is likely to be correlated with the determination of peptide-binding specificity (see below).

The phosphotyrosine-binding site

The first view of SH2-phosphotyrosine interactions was provided by the crystal structures of Src SH2 bound to two non-specific pentapeptides, Y*VPML (SEQ ID NO:2)

and Y*LRVA (SEQ ID NO:3), both with amino-terminal phosphotyrosine (Y*). These structures were determined before a high affinity peptide was identified for the Src SH2 domain. Although these structures do not explain the basis for peptide specificity, they capture the essential elements of the phosphotyrosine-recognition mechanism, as confirmed by the subsequent determination of the structures of high-affinity complexes (see below). The Y*VPML structure was determined at 1.5 Å resolution, and enabled a rather precise determination of the interactions of the phosphotyrosine with the SH2 domain. Three positively charged residues (ArgαA2, ArgßB5 and LysßD6) interact with the phosphotyrosine. In addition, residues from the BC loop contribute a number of hydrogen bonds. Very similar interactions with phosphotyrosine are seen in the structures of the Src and Lck SH2 domains bound to a high-affinity peptide [Eck et al (1993) *Nature* 362:87-91; Wakeman et al (1993) *Cell* 72:779-790]. One difference between the Src and Lck structures is the replacement of CysßC3 by serine in Lck, the oxygen of which provides and additional hydrogen bond to the phosphate.

ArgßB5, which is strictly conserved in all known SH2 domains, is part of the FLVRES (SEQ ID NO:4) sequence which is a signature of SH2 domains [Koch et al (1991)]. The strict requirement for arginine at this position has been demonstrated by mutagenesis experiments on the Abl SH2 domain when lysine substitution led to loss of function [Mayer et al (1992) *Mol. Cell. Biol.* 12:609-618]. This requirement is readily explained as the two terminal nitrogens of the side chain of ArgßB5 form strong hydrogen bonds to two of the oxygen atoms of the phosphate [Wakeman et al (1992)]. This arginine is almost entirely inaccessible to solvent in the structure of the complex. The other two charged residues interact with the phosphotyrosine ring in addition to making contacts with the phosphate or the BC loop.

ArgαA2 forms a strong hydrogen bond between its N atom and a phosphate oxygen and an amino-aromatic interaction between a terminal nitrogen and the ring system thus recognizing simultaneously the phosphate group and the ring system [Wakeman et al (1992)]. The plane of the guanidio group of the arginine side chain is perpendicular to the plane of the benzene ring, i.e., a parallel stacking interaction is not formed. Rather the nitrogen atom of the arginine is 3.5 Å above the center of the aromatic ring, close to the optimal distance suggested by previous analyses of amino-aromatic interactions [Burley et al (1986) *FEBS Lett.* 203:139-143; Perutz et al (1986) *J. Am. Chem. Soc.* 108:1064-1078; Burley et al (1988) *Adv. Protein Chem.* 39:125-189]. These interactions are thought to be comparable in strength to a hydrogen bond between non-charged partners [Levitt et al (1988) *J. Mol. Biol.* 201:751; Loowenthal et al (1993) *J.*

Mol. Biol. 224:759-770]. In the structures of the high-affinity complexes this orientation of the arginine side chain is further stabilized by hydrogen-bonding interactions between both terminal nitrogens and the carbonyl group of the peptide residue preceding the phosphotyrosine [Eck et al (1993); Wakeman et al (1993)].

The third charged residue at the binding site, LysßD6, provides a hydrophobic platform for one face of the aromatic ring. This side chain appears to be somewhat more mobile than the arginines, as judged from the electron density, and takes up slightly different conformations in the Src and Lck structures. In Src, the terminal nitrogen of the lysine interacts with the edge of the aromatic ring and forms a hydrogen bond with ThrßC3, whereas in Lck, it interacts with the phosphate group.

Peptide recognition

The sequence specificities of phosphotyrosine-SH2 interactions were first determined for p85, GAP and PLAγ by mutagenesis experiments and peptide competition assays using the EGF and PDGF receptors [Cantley et al (1991); Escobedo et al (1991) *Mol. Cell. Biol.* 11:1125-1132; Fantl et al (1992) *Cell* 69:413-423; Kashishian et al (1992) *EMBO J.* 11:1373-1381; Rotin et al (1992) *EMBO J.* 11:559-567]. These studies showed that relatively short peptide sequences sufficed to bind to SH2 domains with high affinity. For example, p85α binds preferentially to peptides with methionine or valine at the first position following the phosphotyrosine, and methionine at the third position, with very little selectivity at the second position. A similar sequence specificity has been found for interactions of p85 SH2 domains with the insulin receptor substrate IRS-1 [Shoelson et al (1992) *Proc. Natl. Acad. Sci. USA* 89:2027-2031]. In contrast, GAP binds to a site on PDGF receptor with methionine and proline at the first and third positions, respectively [Kashishian (1992); Kazlauskas et al (1992) *Mol. Cell. Biol.* 12:2534-2544]. A significant observation was that deletions of 12 and 22 residues in regions of the PDGF receptor ~10 residues away from GAP-binding sites did not reduce binding, indicating that the tertiary structure of the receptor in this region is not the primary factor determining binding affinity [Kashishian et al (1992)].

The three residues immediately carboxy-terminal to the phosphotyrosine are apparently the most important in determining the relative binding affinities of phosphopeptide. Based on this, Cantley and co-workers [Songyang et al (1993] used a simple selection method to determine the amino acid preferences at these positions for a number of SH2 domains. Those of Src family tyrosine kinases (including Src, Lck, Fyn and Fgr) select

for the sequence Y*EE (more generally, phosphotyrosine followed by two non-basic polar residues and a large hydrophobic residue). All 11-residue peptide [EPQY*EEIPIYL (SEQ ID NO:5), referred to as the "YEEI" (SEQ ID NO:6) peptide] binds to the Src and Lck SH2 domains with high affinity. The K_d for this peptide binding to glutathione S-transferase fusion proteins incorporation the SH2 domains of Src and Lck is less than 10 nM, measured by surface plasmon resonance [Payne et al (1993) *Proc. Natl. Acad. Sci. USA;* 90:4902-4906]. The structures of the Src and Lck SH2 domains bound to this peptide have been determined at resolutions of 2.7 and 1.8 Å resolution, respectively [Eck et al (1993); Wakeman et al (1993)]. In contacts with the SH2 domain only at the phosphotyrosine, the YEEI (SEQ ID NO:6) peptide binds in an extended conformation, with tight anchoring of the phosphotyrosine and the following three residues.

Of particular importance is the identification of a hydrophobic pocket on the surface of the SH2 domain into which the side chain of the isoleucine is inserted. In contrast to the phosphate-binding site, which is highly conserved, the side-chains that line this pocket are quite variable (the ßD, ßE, EF and BG regions). The selection for glutamate at the first and second positions following the phosphotyrosine may arise from the presence of two basic residues (LysßD3 and ArgßD'1) that interact somewhat loosely with these side chains. Several SH2 domains, such as those of p85, select instead for hydrophobic residues at the first site. One factor for this specificity may lie in insertions in the BG loop, which is longer in p85 and may provide additional interactions at this site [Booker et al (1993) *Cell* 73:813-822].

The SH2 domain is a relatively small module for achieving high-affinity interactions with specific peptides. Proteins such as proteases and immunoglobulin domains achieve high specificity by providing grooves or crevices into which the peptide can fit. The relatively small and flat surface of the SH2 domain does not allow the peptide target to be engulfed by the domain when it is bound. Instead, the SH2-peptide interaction resembles, in the case of Src and Lck, a two-pronged plug (the peptide) engaging a two-holed socket (the SH2 domain).

SH3 structure and peptide binding

SH3 domains also have a common architecture and binding site but the details of the structure appear to be more variable than for the SH2 domains. Perhaps because of this, the prediction of the secondary structural elements of this domain [Benner et al

(1993) *J. Mol. Biol.* 229:295-305] have met with somewhat less success.

The overall fold consists of two anti-parallel ß-sheets that pack against each other so that their strands are nearly orthogonal. In one sheet, two adjacent strands (a and e) are formed from residues that are close to the amino and carboxyl termini, which brings the junctions of the domain close together in a manner reminiscent of the SH2 domains. An additional strand (b or b1) is hydrogen bonded to the amino-terminal strand a and completes the sheet. The crystal structure of Fyn, which is closely related in amino acid sequence to Src, does not appear to have the same secondary structure in this area, but the topology is similar. The two solution structures for p85 differ in some details, but again the topology is similar. Slightly different strand pairings, particularly for strand b, and different short helical connectors between the strands are reported for p85. As the experimental conditions and numbers of nuclear Overhauser effect and hydrogen bond restraints used in both structure determinations were very similar, the origin of this difference is obscure.

Conserved residues, many of them hydrophobic, are clustered on one face of the molecule. The overall surface of the SH3 domain is not unusually hydrophobic, but this region, formed by the a/b and b/c loops and the surface associated with the d and e strands is markedly hydrophobic [Musacchio et al (1992) *Nature* 359:851-855]. Experimental evidence from perturbation of chemical shifts rapidly exchanging peptides supports this hypotheses. These experiments were carried out on Src SH3 by measuring the perturbation of ^{2}H and ^{15}N chemical shifts [Yu et al (1992)], and on p85 SH3 by measuring changes in ^{2}H chemical shifts [Booker et al 1993)]. The peptides used in the Src study were based on sequences found in 3BP1, a protein that interacts with the Abl SH3 domain [Chicchetti et al (1992); Ren et al (1993)]. A 21-mer peptide analogous to dynamin believed to be specific for p85αSH3, was used to map out the binding site for that domain [Booker et al (1993)]. The functional importance of SH3 domains has generally been less well characterized than that of SH2 domains, but the situation is rapidly changing (see Ren et al (1993) and references therein).

Src Homology domains are building blocks in many proteins involved in intracellular signal transduction. Detailed understanding of the pathways involving these domains is complicated by the substantial range of individual specificities in SH2 and SH3, and their combination into large proteins that can form multiple homo- and heteromolecular associations. The reductionist approach to studying individual domains has been very successful (Kuriyan et al (1993) *Curr. Op. Stru. Biol.* 3:828-837; Pawson (1995) *Nature*

373:373-80; Cohen et al (1995) *Cell* 80:237-248). Nonetheless, the interactions between the domains are still poorly understood. These interactions are likely to be of significance in explaining more fully the complete activities of the signal-transducing complexes. Therefore, a detailed structural picture of the inter- and intra-molecular organization of the domains is a significant objective. Several cases of multiple SH2 and SH3 domain-containing constructs have been studied structurally such as Lck SH(32) (Eck et al (1994) *Nature* 368:764-69), Grb2 SH(323) (Maignan et al (1995) *Science* 268:291-93), and Abl SH(32), however, there are technical limitations to current structural approaches. In the crystalline state, packing forces may be of the same magnitude as the weak interdomain forces, contributing to the limits associated with the interpretation of diffraction studies. Solution studies by NMR are only applicable in a molecular weight range of less than 30 kDa. For NMR, time-averaged nOe constraints in rapidly exchanging conformations are not readily interpretable. As a complement to direct structural methods, the investigation of such multidomain complexes using "consolidated" ligands is proposed. These ligands, having multiple binding portions, may be expected to bind with high affinity and specificity when a linker between the two affine segments is of the correct length and there is little affinity of the linker itself. The consolidated ligands do not necessarily resemble any natural ligand.

The proposed "consolidated" ligands are similar in concept to affinity reagents, with the modification that the second functionality is a binding element rather than a reactive moiety. Consolidated ligands may also be useful reagents for studies of the cell biology of the signal transduction complexes (e.g., in combination with antibodies or reporter groups including fluorescent tags) and may provide leads into possible classes of diagnostic and therapeutic agents in the many areas of pathology in which SH domains are involved. This approach will be useful where micromolar affinities of ligands to single SH domains provide insufficient affinity, and hence, specificity for pharmacological action (Gibbs et al (1994) *Cell* 77:175-78; Levitzki and Gazit (1995) *Science* 267:1782-88).

The following examples are presented in order to more fully illustrate the preferred embodiments of the invention. They should in no way be construed, however, as limiting the broad scope of the invention.

EXAMPLE 1

The protein target for this study is the "regulatory apparatus", SH(32) (Eck et al (1994) *Nature* 368:764-69), of human Abelson Kinase ("Abl"). Abl was originally isolated as the gene product from *abl* of murine leukemia virus; the human *abl* gene has been isolated and shown, in cases of chronic myelogenous leukemia, to be a causative factor in a fusion following chromosomal translocation [reviewed in Mayer et al (1995) *Current Biology* 5:296-305]. Ligands of moderate affinity have been identified for the isolated SH2 [Zhu et al (1994) *Oncogene* 9:1379-85] and SH3 domains [Cicchetti et al (1992) *Science* 257:803-6; Ren et al (1993) *Science* 259:1157-61]. Very much higher affinities for some ligand/SH2 interactions have been reported for other SH2s using surface plasmon resonance. This apparent affinity may result from an avidity effect of the use of GST-fusion proteins [Ladbury et al (1995) *Proc. Natl. Acad. Sci. USA* 92:3199-3203]. For Abl SH3, a crystal structure of the complex of the ligand 3BP-1 is available [Musacchio et al (1994) *Nature Structural Biology* 1:546-51], indicating that the ligand is in the so-called class I orientation [reviewed in Pawson (1995) *Nature* 373:373-80]. Using the general positioning of ligands to SH2 as a framework [Kuriyan et al (1993) *Curr. Op. Stru. Biol.* 3:828-37], the likely orientation of the "2BP-1" ligand PVY*ENV amide [Zhu et al (1994) *Oncogene* 9:1379-85] was modelled. In an SH(32) model (Figure 1), these ligands are closest at their C-termini, and there is more than 20 Å separating them and about 32 Å separate the C-terminus of the SH3 ligand from the N-terminus of the SH2 ligand. An initial design of a "consolidated" ligand for both domains uses this model. This Example relates to a consolidated ligand having two affine ligands joined together. The idea may be extended to as many as are synthetically practical, and these need not be linearly arranged.

Preparation and synthesis of ligands. Solid-phase syntheses of branched sequences, as well as linear controls, were carried out either manually or on a continuous-flow Millipore 9050 instrument, starting with Fmoc-PAL-PEG-PS resins (0.6 g scale, loading 0.16 mM/g ("normal") or 0.49 mM/g ("high load")). Appropriate N^{α}-Fmoc-amino acids (4.0 equiv.) were used with Fmoc removal by piperidine-DMF(1:4, 2 x 8 min), and standard BOP/HOBt/NMM activation chemistry described by Alberico et al [(1990) *J. Org. Chem.* 55:3730-3743 and references therein). Side-chain protection was provided by a tert-butyl (tBu) ether and ester for threonine and glutamic acid amino acids respectively, while the N-terminal proline residues (both branches) were protected by Boc. Syntheses of branched ligands began with incorporation onto PAL of the C-terminal lysine as its N^{α}-Fmoc, N^{ϵ}-Dde derivatives (Bycroft et al (1993) *Chem. Soc. Chem.*

Commun. 1993:778-779). First, the 3BP chain was assembled off the N^{ϵ}-amino branch, ending with Boc-Pro. Then the Dde group was removed by hydrazine hydrate-DMF (1:49; 6 x 5 min), and assembly of the oligoglycyl spacer and 2BP sequence continued off the N^{ϵ}-amino side-chain. NMM was substituted by DIEA at the step when N^{ϵ}-Fmoc-phospho-L-tyrosine was coupled to the peptide-resin, and for the subsequent couplings (Ottinger et al (1993) *Biochemistry* 32:4354-61). In an alternative method, an N^{α}-Fmoc-Tyr-OH with an unprotected phenolic side-chain was incorporated by standard BOP/HOBt/NMM protocol. After completion of the chain assembly on-resin phosphorylation was performed with di-tert-butyl N,N-diethylphosphoramidite/IHA-tetrazole, followed by oxidation using anhydrous tert-butyl hydroperoxide. An H-phosphonate control (VIII) peptide was made by omitting this last step. Final cleavage and deprotection reactions were carried out using Reagent K (TFA-phenol-H_2O-thioanisole-1,2-ethanedithiol (35:2:2:2:1)), and the crude products were purified by semi-preparative reversed-phase high-pressure liquid chromatography (RP-HPLC) to provide peptide products as white powders after lyophilization; overall yields of 20-40%, purities >93% by analytical HPLC. Amino acid analyses gave consistent ratios, and the products were characterized further by mass spectra.

Consolidated ligand I (Table 1) then joins the C-termini of the individual 2BP and 3BP ligands using a branch through the side-chain of a lysine residue. By model building, the methylene segments of the lysyl side-chain and the oligoglycyl linker are expected to provide sufficient separation between the C-termini of the individual ligands. A series of consolidated ligands, with different ligand segments for SH3, different lengths and orientations of the linker, and different analogs of the phosphotyrosyl residue in the ligand for SH2, were synthesized and tested for affinity to Abl SH(32), and in some cases to SH3 and SH2.

This design and synthesis of consolidated ligands is shown to be effective in Table 1. Approximately two-orders of magnitude increase in affinity was observed compared to unbranched equivalents (I vs II, IV vs. V). An order of magnitude increase is observed by comparing the most strongly bound single ligand to the equivalent consolidated

ligand (2BP1 vs. IV). The affinity of a consolidated **TABLE I**

LIGAND	STRUCTURE	AFFINITY(nm) SH(32)	SH2	SH3
2BP-1	PVY*ENVam (SEQ ID NO:7)	2,350(340)	2,020(230)	
3BP-1	PPTMPPPLPP (SEQ ID NO:8)	56,000(900)		
3BP-2	PPAYPPPPVP (SEQ ID NO:9)	10,700(300)		10,500 (200)
I	(2BP-1)-G_6] (3BP-1)Kam (SEQ ID NO:10)	487(16)		
II	(2BP-1)G_6K(3BP-1) (SEQ ID NO:11)	18,000 (1,200)		
III	(2BP-1)G_5] (3BP-2)Kam (SEQ ID NO:12)	503(7)		
IV	(2BP-1)G_6] (3BP-2)Kam (SEQ ID NO:13)	249(5)	14,200 (900)	5,370 (340)
V	(2BP-1)G_6K(3BP-2)am (SEQ ID NO:14)	22,100 (1,700)		
VI	(2BP-1)G_7] (3BP-2)Kam (SEQ ID NO:15)	327(6)		
VII	(2BP-1)G_8] (3BP-2)Kam (SEQ ID NO:16)	389(23)		
VIII	H-phosphonate analog of IV	1,810(80)		
IX	desphospho-analog of IV	13,800(920)		6,050 (82)

Legend to Table I. Recombinant Abl SH domains were obtained by expression of GST fusions, which were the generous gift of B. J. Mayer and D. Baltimore, and domains were cleaved and purified as described previously [Waksman et al (1993) *Cell* 12:779-90]. Concentrations of proteins and peptides were determined by UV absorbance at 264 nm, using values of 1,752 for phosphotyrosyl, and 840 for tyrosyl, for mean residue absorption coefficients [Waksman et al (1993) *Cell* 12:779-90]. Affinities were measured by intrinsic fluorescence quenching using a Perkin-Elmer 760-40 fluorimeter, with excitation at 290 nm, and emission at 345 nm, with effective band widths of 2 and 17 nm respectively. The sample compartment was maintained at 18°C, and a mini-magnetic stirrer was used for continuous mixing. Titration of SH domains was accomplished by making appropriate sequential additions from the peptide stock solution to the domain(s) in a 1-cm square quartz cell. Fluorescence was recorded at two minutes after each addition. 1 mL of protein solution (140 mM NaCl, 5 nM phosphate, pH 7.3) with 1 mM DTT was used. Protein concentrations were approximately 500 nM. Data were analyzed using the program Origin (Microcal Software, Inc.) to fit the equation:

$$[\text{total ligand}] = (K_d*(F-F_o) + (F_{max}-(F-F_o))*(F-F_o)*[\text{total protein}]/F/_{max})/(F/_{max}-(F-F^o))$$

where F is the observed fluorescence quenching at ligand concentration [[total ligand], F_{max} is the quenching of protein saturated with ligand, and F_o is the fluorescence without ligand. K_d, F_{max}, and F_o are treated, as fitted

parameters. Symbol "]" implies N$^{\epsilon}$-lysine side-chain ('Branched'). "Am" = C-terminal carboxamide. Y* = O-phospho-L-tyrosine. The "Standard Deviation", SD, used was that calculated using Origin for a single titration. Replicate titrations, were done, showed a standard deviation about twice the single set SD. The absolute accuracy of these values depends critically on the determination of concentrations of reagents. Nonetheless the *relative* values reported here are highly precise, for the purposes of the comparisons drawn. Note that ligand IV was prepared by two different routes: 1) direct solid phase synthesis incorporating phosphotyrosine, and 2) by post-assembly phosphorylation; their affinities were identical within experimental error.

ligand changes when a subligand is modified proportionally to the change of affinity of separate ligand (**I/3BP-1** vs. **IV/3BP-2**). The simple linker chemistry used does not interfere with the subligands affinities to the individual SH3 or SH2 domains. Experiments with the single ligands as inhibitors show classical competitive binding. The expectation of increased affinity from simple physical chemical concepts is somewhat higher. It may be assumed that the principal method by which apparent affinity increases using this approach is the reduction in degrees of translational freedom, when one subligand is bound [Jencks (1981) *Proc. Natl. Acad. Sci. USA* **78**:4046-4050]. For perfect linker length and geometry, a consolidated ligand affinity up to the product of the individual association constants is predicted. Increases in degrees of rotational freedom associated with each glycyl residue and with the lysyl methylenes reduce this affinity considerably. Calorimetric studies of the complex formation are used to test this simple model. The values of affinities of ligands with variable linker lengths (III-G_5, IV-G_6, VI-G_7, VII-G_8) are most readily interpreted as indicating an optimal length of about G_6, with little interaction between the linker and the SH(32) protein. Analogs of the 2BP ligand of **IV**, viz. the H-phosphonate (**VIII**), and the pY→Y of **IV** (**IX**) lead to decreases in affinity. The simplest model of how the consolidated ligand binds is that there are specific contacts at the subligand sites and a solvated linker region. Using a model derived from NMR studies of the SH3 and SH2 domains, assuming that the orientation for Abl is similar to that seen in Lck [Eck et al (1994) *Nature* **368**:764-69], the two subligands are correctly accommodated, and no specific interactions are forced on the linker segment (Figure 1). The specificity of interaction is much increased also (Table 1), so that **IV** will now discriminate between the complete SH(32) and its subdomains by more than an order of magnitude.

There are many possible applications of these consolidated ligands and their analogs. The examples in Table I show that for Abl SH(32), the two ligand sites can be oriented as shown in Figure 1, that the 3BP-1 and 3BP-2/ligands for SH3 (in I and IV) are bound in the same direction, and that the SH2 and SH3 binding sites do not interfere with each other. Applications are expected for utilization of other SH(32)-containing proteins for similar mapping purposes, or for obtaining high affinity and specificity of binding for investigational or therapeutic purposes [Brugge (1993) *Science* **260**:918-19]. Some examples for utilization of SH(32) include dual and higher-subligand consolidations for the adaptor protein Grb-2 and reagents like those shown here for Abl, though possibly linear rather than branched [Feng et al (1994) *Science* **266**:1241-1247] for Src-family tyrosyl kinases [Gilmer et al (1994) *J. Biol. Chem.* **269**:31711-31719]. A major challenge in the investigation of intracellular signal transduction involves the relatively transient nature of the signal transducing complexes formed. Highly specific reagents may permit trapping of or selective interference with these complexes.

The subligands in such consolidated ligands may be extended to other components of signal transducing complexes, for example the active sites of kinases, or binding sites of pleckstrin homology domains. For Abl, the consolidated ligands shown here indicate that ligands may have sufficient affinity and specificity to block the Bcr-Abl fusion kinase that is the predominant pathogenic agent in chronic myelogenous leukemia.

EXAMPLE 2

A consolidated dual SH3 binding peptide was constructed for Grb2. The peptide

P P P V P P R R R G G G G G G R R R P L P P L P P P -amide (SEQ ID NO:17)

had an affinity of 371 ± 33 nM for GST-Grb2 from a commercial source. This peptide uses the type I and type II binding to SH3 domains simultaneously in a

single consolidated ligand to occupy both SH3's of Grb2.

EXAMPLE 3

A consolidated ligand to Src SH(32) was prepared having the following sequence:

N-acetyl-Y^{+}E E I E G G G G G G G G A F A P P L P R R -amide.

The binding of the ligand is tested with the appropriate proteins as described in Example 1.

The chemistry used in these consolidated ligands is relatively simple, and was selected to provide flexible linkers [Maraganore et al (1990) *Biochemistry* **29**:7095-7101]. Linkers of the correct size and rigidity provide greater affinities. The structure of the Abl SH(32)/consolidated-ligand complex permits rational design of linkers of an optimal rigidity and size, and also permit identification of additional interactions for increased affinity. Peptidomimetic, and non-peptidic linkers are generally practical, as are other subligands including lipids, steroids, carbohydrates, and nucleic acids. These consolidated ligands present novel opportunities for chemistry at the interface with biology.

While the invention has been described and illustrated herein by references to various specific material, procedures and examples, it is understood that the invention is not restricted to the particular material combinations of material, and procedures selected for that purpose. Numerous variations of such details can be implied as will be appreciated by those skilled in the art.

WHAT IS CLAIMED IS:

1. A consolidated ligand comprising ligands to two or more protein domains, wherein a first ligand to a first domain of a target protein is connected via a linker to a ligand to a second domain of a target protein, wherein the consolidated ligand has a greater affinity and/or specificity for the first and second domains than either the ligand for the first domain or the ligand for the second domain alone.

2. The consolidated ligand of Claim 1, wherein the first and second domains are from the same target protein.

3. The consolidated ligand of Claim 1, wherein the first and second domains are from different target proteins.

4. The consolidated ligand of Claim 1, wherein the ligands comprise peptides.

5. The consolidated ligand of Claim 4, wherein the ligand has an amino acid sequence selected from the group consisting of SEQ ID NOS:10-17.

6. The consolidated ligand of Claim 1 labeled with a detectable label.

7. The consolidated ligand of Claim 6 wherein the label is selected from enzymes, chemicals which fluoresce and radioactive elements.

8. The consolidated ligand of Claim 1, wherein the target protein is a protein involved in signal transduction.

9. The consolidated ligand of Claim 8, wherein the protein involved in signal transduction is a protein tyrosine kinase.

10. The consolidated ligand of Claim 9, wherein the first and seconds domains comprise a Src homology 2 domain or a Src homology 3 domain.

11. The consolidated ligand of Claim 10, wherein the target protein is Abl, APRF, Arg, Bcl-2, Blk, Btk, c-erbB-2, CAK, cbl, Cbl-B, CCK-2, CD25, CDK, CFSR, clk, crk, crkl, E2A, EGFR, ena, erbB2, ERK-2, Esk, FAK, Fes, Fgr, Fit3, Flk2, FLT1, fos, Fps, FUS3, Fyn, gap, grb2, grb7, HOX11, HER2, HGF/SF, IGR-IR, IGFR, int-2, iyk, Jak3, jun, KDR, KIT, Lck, Lyl1, lyn, MAP kinase, Matk, MCK-10, MET, MO15, mta1, MUC1, nck, nek, NEU, p53, PDGFR, PI3K, PIM, PLC-gamma, PTB1b, PTP1C, PTPN6, RAF, Rak, ras, RB, Ret, Rlk, ROS, SCL, Shc, Sos, SRC, STK1, SYK, Syp, Tec, tie, TRK, tsk, Ttg1, Ttg2, Txk, Vav, or yes.

12. The consolidated ligand of Claim 8, wherein said signal transduction originates with binding of a receptor to EGF, EPO, FGF, IGF, IL, int2, LIF, PDGF, SCF, TGF, TNF or VEGF.

13. A pharmaceutical composition comprising the consolidated ligand of Claim 1 in combination with a pharmaceutically acceptable carrier.

14. A nucleic acid molecule encoding the consolidated ligand of Claim 1.

15. The nucleic acid molecule of Claim 14, comprising:

(A) a DNA sequences that code on expression for an amino acid sequence selected from the group of SEQ ID NOS:10-17;

(B) DNA sequences that hybridize to any of the foregoing DNA sequences under standard hybridization conditions.

16. The nucleic acid molecule sequence of Claim 14 which is operably linked in proper reading frame to sequences which direct the expression of said nucleotide sequence.

17. A host cell transformed with the nucleic acid molecule of Claim 16.

18. The host cell of Claim 17 wherein the unicellular host is selected from the group consisting of *E. coli*, *Pseudomonas*, *Bacillus*, *Streptomyces*, yeasts, CHO, R1.1, B-W, L-M, COS 1, COS 7, BSC1, BSC40, and BMT10 cells, plant cells, insect cells, and human cells in tissue culture.

19. The nucleic acid molecule of Claim 16, wherein said sequences which direct expression are selected from the group consisting of the early or late promoters of SV40 or adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase, the promoters of acid phosphatase and the promoters of the yeast α-mating factors.

20. An antibody to the consolidated ligand of Claim 1.

21. The antibody of Claim 20, wherein the antibody is polyclonal.

22. The antibody of Claim 21, wherein the antibody is monoclonal.

23. An immortal cell line that produces a monoclonal antibody according to Claim 22.

24. The antibody of Claim 20 labeled with a detectable label.

25. The antibody of Claim 24 wherein the label is selected from enzymes, chemicals which fluoresce and radioactive elements.

26. A method for detecting the presence or activity of a target protein, wherein said target protein is measured by:

A. contacting a biological sample from in which the presence or activity of said target protein is suspected with a consolidated ligand to said target protein under conditions that allow binding of said consolidated ligand to said target protein to occur; and

B. detecting whether binding has occurred between said target protein from said sample and the consolidated ligand;

wherein the detection of binding indicates that presence or activity of said target protein in said sample.

27. A method for detecting the presence and activity of a polypeptide ligand associated with a cellular condition comprising detecting the presence or activity of a target protein according to the method of Claim 26, wherein detection of the presence or activity of the target protein indicates the presence and activity of a polypeptide ligand associated with a given cellular condition.

28. The method of Claim 27 wherein said cellular condition is agammaglobulinemia, AIDS, ALL, angiogenesis, breast cancer, carcinoma, chromic myelogenous leukemia, colon carcinoma, colorectal cancer, diabetes, erythroleukemia, gastric cancer, hematopoiesis, Kaposi's Sarcoma, leukemia, liver regeneration, Lyme disease, megakaryocytopoiesis, melanoma, neuroblastoma, organogenesis, osteopetrosis, ovarian hyperstimulation syndrome, placental development, severe combined immunodeficiency, ulcerative colitis or Wilms tumor.

29. A method for detecting the binding sites for a target protein to a consolidated ligand, comprising:

A. placing a labeled consolidated ligand sample in contact with a biological sample in which binding sites for said consolidated ligand are suspected;

B. examining said biological sample in binding studies for the presence of said labeled consolidated ligand;

wherein the presence of said labeled consolidated ligand indicates a binding site for said consolidated ligand on said target protein.

30. A method of testing the ability of a drug or other entity to modulate the activity of a target protein, which comprises:
A. culturing a colony of test cells which has a binding site for a consolidated ligand in a growth medium containing the consolidated ligand;
B. adding the drug under test; and
C. measuring the reactivity of said consolidated ligand with the target protein on said colony of test cells.

31. A test kit for the demonstration of a target protein in a eukaryotic cellular sample, comprising:
A. a predetermined amount of a detectably labelled consolidated ligand;
B. other reagents; and
C. directions for use of said kit.

32. A test kit for demonstrating the presence of a target protein in a eukaryotic cellular sample, comprising:
A. a predetermined amount of a target protein;
B. a predetermined amount of a consolidated ligand for said target protein;
C. other reagents; and
D. directions for use of said kit;
wherein either said target protein or said consolidated ligand are detectably labelled.

33. A method of preventing and/or treating cellular debilitations, derangements and/or dysfunctions and/or other disease states, comprising administering to a recipient a therapeutically effective amount of the consolidated ligand of Claim 1.

34. The consolidated ligand of Claim 1, wherein the linker is a peptidomimetic.

35. The consolidated ligand of Claim 1, wherein the linker is non-peptidic.

36. The consolidated ligand of Claim 35, wherein the linker comprises a lipid, steroid, carbohydrate, or nucleic acid.

37. The consolidated ligand of Claim 1, wherein the linker is 10-50 A in length.

38. The consolidated ligand of Claim 37, wherein the linker is 20-40 A in length.

39. The consolidated ligand of Claim 38, wherein the linker is 30-35 A in length.

40. The consolidated ligand of Claim 1, wherein the linker comprises glycyl moieties.

41. The consolidated ligand of Claim 40, wherein the linker comprises 2 to 20 glycyls.

42. The consolidated ligand of Claim 41, wherein the linker comprises 3 to 12 glycyls.

43. The consolidated ligand of Claim 42, wherein the linker comprises 5 to 8 glycyls.